The use of infrared thermography to investigate emotions in common marmosets

Ermatinger, F A ; Brügger, Rahel Katharina ; Burkart, J M

DOI: https://doi.org/10.1016/j.physbeh.2019.112672

The use of infrared thermography to investigate emotions in common marmosets

Ermatinger, F.A., Brügger, R.K. & Burkart, J.M.

Department of Anthropology
University of Zurich
Winterthurerstrasse 190
8057 Zurich
Switzerland

Corresponding Author
Florine Anna Ermatinger
florine.ermatinger@gmail.com
Winterthurerstrasse 190
8057 Zurich
Switzerland

Further information

- Every figure should be in color

Abstract

Measuring body surface temperature changes with infrared thermography has recently been put forward as a non-invasive alternative measure of physiological correlates of emotional reactions. In particular, the nasal region seems to be highly sensitive to emotional reactions. Several studies suggest that nasal temperature is negatively correlated with the level of arousal in humans and other primates, but some studies provide inconsistent results. Our goal was to establish the use of infrared thermography to quantify emotional reactions in common marmosets (Callithrix jacchus), with a focus on the nasal region. To do so we exposed 17 common marmosets to a set of positive, negative and control stimuli (positive: preferred food, playback of food calls; negative: playback of aggressive vocalizations, teasing; control: no stimulus). We compared nasal temperature before and after the stimuli and expected that highly aroused emotional states would lead to a drop in nasal temperature. To validate the thermography measure, we coded piloerection of the tail as an independent measure of arousal and expected a negative correlation between the two measures. Finally, we coded physical activity to exclude its potential confounding impact on nasal temperature. Our results show that all predictions were met: the animals showed a strong decrease in nasal temperature after the presentation of negative arousing stimuli (teasing, playback of aggressive vocalizations). Furthermore, these changes in nasal temperature were correlated with piloerection of the tail and could not be explained by changes in physical activity. In the positive and the control conditions, we found systematic sex differences: in males, the preferred food, the playbacks of food calls, as well as the control stimulus led to an increase in nasal temperature, whereas in females the temperature remained stable (preferred food, control) or decreased (playback of food calls). Based on naturalistic observations that document higher food motivation and competition among female marmosets, as well as stronger reactions to separation from group members in male marmosets, these sex differences corroborate a negative correlation between arousal and nasal temperature. Overall, our results support that measuring nasal temperature by infrared thermography is a promising method to quantify emotional arousal in common marmosets in a fully non-invasive and highly objective way.

Keywords

Infrared thermography, Nasal skin temperature, Common marmosets, Emotions, Autonomic reaction, Piloerection, Arousal

1 Introduction

Identifying emotional states remains one of the big challenges in animal behavior research since in contrast to humans, animals cannot document their subjective experience in self-reports. However, in addition to subjective experiences, emotional states also correlate with behavioral and physiological responses that are accessible to empirical investigation in animals [1–3].

Physiological markers are particularly promising because they can be quantified objectively. Unfortunately, techniques to measure them tend to be invasive and restrict the movement of the subject. For example, EEG or fMRI can be used to measure brain activity patterns but require attaching electrodes or placing the subject in an fMRI tube [4]; changes in heart-rate can be assessed by electrocardiography that requires electrodes on the subjects body [5–7]; quantifying the galvanic skin response necessitates a finger transducer [8] and blood pressure is measured by a cuff or an arterial catheter [9,10].

The invasiveness of such methods limits their use to measure emotions, especially in animals. First, the measurement interferes with the behavior and is likely to produce a strong emotional reaction in itself. Furthermore, habituating the subject to the procedures is often time intensive or even impossible [1,6,11]. A fundamental limitation in the study of animal emotions is thus the scarcity of non-invasive methods to measure physiological reactions. This limitation is particularly striking when the goal is to quantify emotional arousal during naturalistic conditions or social interactions, which is an increasingly important domain of research [3,12].

In recent years, infrared thermography has been pioneered as a fully non-invasive, contact-free technique to examine emotional states. This technique measures emotion-related surface temperature changes [13]. Several studies have focused on temperature changes in highly exposed body regions such as the nose, fingers, paws or the tail of animals [3,6,14–16]. The temperature of these areas seems to be highly sensitive to emotional changes. However, results are sometimes incongruent, and it is not yet clear which temperature reactions correspond to what kind of emotion and how sensitive such measures are.

In humans, first studies showed a trend to a decrease in temperature in peripheral body regions during negative, high-arousal states. Thus, a decrease in nasal or finger temperature has for example been observed in response to stimuli inducing mental stress [17], fear [16], guilt [14] or when experiencing empathy for a child in distress [18]. These temperature drops in peripheral body regions are mostly attributed to changes in subcutaneous perfusion that are controlled by the sympathetic branch (SNS) of the autonomic nervous system. The idea is that during the “fight or flight” response, the SNS induces peripheral vasoconstriction and hence a decrease in blood flow and cooling of the respective regions [13,14,19]. This mechanism is most likely caused by the immediate norepinephrine release by the sympathetic nervous system, that occurs within seconds after the presentation of stressful stimuli. Minutes later, especially with strong negative stimuli, the thermal signature could further change due the secretion of cortisol by the adrenal gland [20,21].

Peripheral temperature drops, however, appear to accompany arousal in general, and not only in combination with negative valence. In fact, various studies that examined thermal reactions to positive stimuli suggest that the respective impact of arousal on body temperature is independent of whether the emotion is of positive or negative valence. For instance, Kosonogov et al. [22] measured nasal temperature while human participants were watching standardized emotional pictures. They could show that arousal, but not valence, of the induced

psychological state had an effect on nasal temperature. Furthermore, Salazar-Lopez et al. [23] found that nasal temperature decreased during positive empathic reactions (empathic happiness). Another study found that playing and laughing, behaviors most likely linked to positive, high arousal emotions, evoked nasal temperature drops in human children [24]. Consistent with this general effect of arousal, positive, relaxing situations were often linked to the opposite thermal reaction; an increase in peripheral body temperature. For example, human participants instructed to relax [25] or watching positive, relaxing video clips [26] showed an increase in finger temperature. Likewise, watching at and thinking about a beloved person evoked an increase in nasal and hand temperature in most and half of the participants, respectively [23].

Whereas the above findings match the suggested pattern of a negative correlation between arousal and the temperature of peripheral body regions in humans, a few studies have shown other reactions. For instance, sexual arousal was linked to an increase in nasal temperature although it is most likely not associated with a relaxed emotional state [27,28]. Further, Salazar-López et al. [23] conducted an experiment in which human participants were exposed to IAPS images (International Affective Picture System: Lang et al. [29]) which elicited psychological states of varying valence and arousal levels. Positive images, both low and high arousal were linked to an increase in nasal temperature while for negative conditions, nasal temperature decreased in response to low arousal but increased in the case of high arousal.

In thermography studies with animals, most results support a general pattern of a negative relationship between arousal and the temperature of peripheral body regions. For example, macaques showed a decline in nasal temperature in response to a threatening person [19] or visual and vocal signals of aggressive conspecifics [8]. Kano et al. [6] found that the nose of chimpanzees became colder in response to playback calls and videos of fighting conspecifics. They validated their findings with behavioral but also additional physiological measurements. Most notably for one of the subjects they reported a decrease of the relative power of the high-frequency band of the heart rate variability spectrum, which is reported to be a reliable indicator of a highly aroused state. A further study [3] investigated thermal reactions of wild chimpanzees in response to naturally occurring species-specific vocalizations. They found that nasal temperature dropped significantly after aversive vocalizations while it stayed about the same after neutral ones. Intriguingly, they also report a significant increase in ear temperature in the neutral conditions but not in response to aversive vocalizations. Potentially underlying physiological mechanisms for these changes in ear temperature are not known. In fact, a different pattern was observed in rabbits and dogs where ear temperature decreased during stress or fear related situations [30,31]. Additionally, stress-related situations evoked a temperature drop in the tail and paws in rats [15].

In contrast to negative stimuli, much less research has been conducted to quantify thermal reactions to positive stimuli in animals. However, a macaque showed an increase in nasal temperature while it was stroked by a human, and the decrease in its heart rate suggested that the monkey became more relaxed during this procedure [32]. In addition, Chotard et al. [33] found a trend for a decrease in nasal temperature in non-human primates when the subjects were exposed to a toy, what was assumed to be positively arousing.

Another macaque study compared nasal temperature among three conditions that were assumed to be negatively or positively arousing: feeding with potential food competition, teasing with food, and playing with a toy. Nasal temperature was only significantly lower in

the feeding condition compared to the other two conditions, which did not show a significant difference in nasal temperature among each other [11]. However, the conditions were not compared to a "neutral" control or a baseline condition and no behavioral markers of arousal were coded, which hampers the interpretation about the link between arousal and nasal temperature. Table 1 summarizes the presented studies on humans and animals, classifying the stimuli according to the suggested valence and arousal level as well as the corresponding thermal reaction.

Table 1 | Summary of studies that examined the link between emotional states and thermal reactions in peripheral body regions (nose, finger, tail or ears). We classified the emotions/psychological states according to the suggested valence and level of arousal as well as the corresponding evoked peripheral temperature change. Some studies are mentioned several times as they examined more than one emotion. The subscript letters indicate whether the emotion was examined in humans (H), non-human primates (P) or other animals (A).

		Arousal	
		Increase	**Decrease**
Valence	**Positive**	**Temperature drop** - Empathic happiness [23]^{H} - Joy [24]^{H} - Playfulness [33]^{P} **Temperature rise** - Low and high, positive arousal due to exposure to IAPS images [23]^{H} - Sexual arousal [27,28]^{H}	**Temperature drop** - No studies **Temperature rise** - Relaxing [25,26,32]^{H,P} - Feeling love [23]^{H}
	Negative	**Temperature drop** - Fear [16]^{H} - Guilt [14]^{H} - Negative empathy [18]^{H} - Unpleasant arousal [22]^{H} - Low negative arousal due to exposure to IAPS images [23]^{H} - Stress, fear [3,6,8,11,15,19,30,31]$^{P, A}$ - Feeding with food competition (compared to teasing and playing) [11]^{P} **Temperature rise** - High negative arousal due to exposure to IAPS images [23]^{H} - Teasing and Playing with a toy (compared to feeding with food competition) [11]^{P}	No studies
	Neutral	**Temperature drop** - Mental Stress [17]^{H} **Temperature rise** - No studies	No studies

In contrast to humans, there are almost no results that contradict the proposed link between arousal and a decrease in peripheral body temperature in animals. However, this could be a simple artefact of the bias for studies about negative, high arousal emotions in animals. Therefore, in order to reliably apply this method to assess the level of arousal in animal studies, it is mandatory to investigate a variety of possible emotions that vary with both regard to valence (positive/negative) and arousal.

A final issue in thermography studies, especially with animals, concerns the role of physical activity [6,11,31]. An alternative explanation for thermal changes that accompany arousal may be that higher arousal simply leads to higher physical activity that in turn influences body temperature. Therefore, to verify the link between emotions and changes in peripheral body temperature, it is important to examine this potentially confounding effect.

In this study we investigated the use of infrared thermography to infer emotional states in common marmoset, a New World monkey species. Marmosets are increasingly used as

model species in social neurosciences and evolutionary anthropology [12,34–39] because of their high degree of sociality, cooperation and allomaternal care, which show striking similarities with humans [37,38]. Marmosets are small, highly active animals and do hardly ever stand still. This makes it difficult to reliably track the temperature of a respective region of interest, especially if its temperature is not very different from the surrounding area. However, the nasal region is prominently visible as a cold spot in the face and hence can be tracked reliably. We therefore decided to focus on nasal temperature as a first investigation about the use of infrared thermography in marmosets.

We experimentally examined nasal temperature changes in 17 common marmosets in five conditions that were supposed to induce psychological states of varying levels of arousal and valence: teasing with food, playback of aggressive vocalizations, preferred food, playback of food calls and control (no stimulus).

Based on previous studies, we expected a negative correlation between arousal and nasal temperature. We a priori assumed that teasing with food and the playback of aggressive vocalizations would evoke a negative, intermediate to high arousal state in the animals, whereas eating a preferred food item was assumed to be a positive, relaxing event. A similar reaction could be expected in response to playbacks of food calls given by group members because previous studies have shown that vocalization of familiar individuals can lead to a stress reduction (decrease in cortisol level) [40]. However, food calls mostly refer to the presence of preferred food items and encourage the animals to search for the food source [41,42]. Therefore, we expected that the food calls could also lead to a positive, but aroused rather than relaxed state. During the control condition, the experimenter turned away and no stimulus was given. This was assumed to be neither positive nor negative and thus not lead to a change in arousal.

Simultaneously to the presentation of the stimuli and the assessment of changes in nasal temperature, we quantified a behavioral marker of arousal throughout the experimental conditions (i.e. piloerection of the tail). This behavioral marker allowed us to validate the findings of the thermal reactions. Finally, we also quantified physical activity, to disentangle whether thermal reactions indeed capture emotional states rather than a mere consequence of physical activity.

2 Methods

2.1 Subjects and housing

We tested 17 adult common marmosets (*Callithrix jacchus*) (see supplementary material S.1 for details about the subjects). They were all born and reared in captive family groups. The animals lived in different groups or pairs that were housed in adjacent enclosures and had only visual and olfactory contact with other groups.

The monkeys were housed in heated indoor enclosures that contained sleeping boxes, heated resting platforms as well as climbing and playing material like ropes and branches. The floor was covered with bark mulch. When weather conditions were suitable, the animals also had access to outdoor enclosures that could be reached via transparent tubes.

The marmosets were fed at least twice a day (porridge with vitamins at 8:30; fresh fruits and vegetables between 11:00 and 12:00) and water was continuously available. During the week they received an additional snack (mealworms, nuts etc.) in the afternoon. For this study the animals were neither food nor water deprived. The monkeys had already been involved in other behavioral experiments but had no previous experience with the current experimenter or a thermography camera.

2.2 Ethics statement

The experiments were approved by the Kantonales Veterinäramt des Kantons Zurich, Switzerland (license number 223/16) and classified as degree of severity = 0.

2.3 Experimental procedure

During the experiments the monkeys were separated in a testing compartment (60*50*50 cm) of a larger testing enclosure in a testing room. This helped to control for spatial variation in environmental conditions (light, temperature) and to reduce the monkey's ability to move around and interact with group members. The monkeys voluntarily entered the testing enclosure from their home cage via a tube system that could be opened and closed by the experimenter. Thus, no direct handling of the animals was required. At the front mesh of the testing enclosure we attached a perch that served as preferred sitting area for the monkeys. Two cameras were placed in front of the testing compartment, a thermography camera (FLIR, T620) and a camcorder (SONY, HDR-CX730E). Prior to the experiments, the animals were habituated to the testing enclosure over a period of several weeks, first with their family group, later alone. Subjects who failed to habituate were excluded from the experiment.

Before an experimental session started, the experimenter opened the entrance to the testing enclosure and allowed free access to all animals of the social group for about two minutes. After this settling-in period the experimenter separated the subject in the testing compartment and lured the other group members out of the testing room. The experimental session started when the subject was quietly sitting on the perch or clinging to the front mesh and did not show obvious signs of stress, such as escape attempts, excessive vocalization or running quickly back and forth.

Each experimental session was divided into a baseline, a stimulation and a post-stimulation phase. The baseline phase lasted for two to three minutes. During this phase the experimenter offered a mealworm or a small piece of nut cookie to the subject periodically (every 20 seconds) through the mesh with a tweezer. This procedure occupied the subject and kept its attention towards the direction of the camera. Although a potential effect of this feeding procedure on the emotional state could not be fully excluded, we expected it to be negligible, especially compared to the other stimuli, as it was only a small amount of food and the monkeys were already habituated to this procedure. More importantly, the procedure was identical for all conditions and could therefore not influence the main results, i.e. differential reactions depending on stimulus type during the stimulation phase.

The stimulation phase started 20 seconds after the last mealworm was handed over in the baseline. In this phase the subject was not offered food but was instead exposed to one of the five stimuli. The post-stimulation phase started right after the stimulus has ended. It was identical to the baseline and lasted for about three minutes.

2.4 Stimuli

We used 5 stimuli that were presented to the subjects in randomized order. Each subject was tested maximally once a day in one of the five conditions. To elicit an increase in negative arousal, we used teasing and a playback of aggressive vocalizations. To evoke positive emotions we used offering highly preferred food and a playback of food calls of familiar individuals. During the control stimulus the experimenter simply stopped interacting with the subject.

During the ***teasing*** condition, the experimenter showed the subject a defrosted cricket *(Acheta domesticus)* or a nut cookie, depending on what the respective subject preferred. The experimenter held the food item right in front of the mesh and pulled it away each time the subject tried to grab it. The teasing lasted maximally 60 seconds. However, the teasing was finished earlier when the subject showed increasing signs of stress (e.g. escape attempts, running back and forth).

The ***playback of aggressive vocalizations (PB-)*** consisted of aggressive calls of common marmosets that were given during intergroup visual contact. The calls were recorded with the Avisoft UltraSoundGate 116H laboratory software and a condenser microphone CM16/CMPA. For all the sessions we used the same recording which did not include calls from group members of any of the subjects. This stimulus lasted maximally 60 seconds but again it was terminated prematurely if the subject showed increasing signs of stress. The playback was broadcast using “iMovie” (Apple, Version 10.1.9) from a laptop (MacBook Pro, Apple) placed in the right or left neighboring compartment of the testing compartment.

In the ***preferred food (food)*** condition, a defrosted cricket *(Acheta domesticus)* was given to the subject by the experimenter with a tweezer through the mesh. The subject was allowed to eat the cricket in its entirety without any disturbance. During the feeding time, the experimenter turned to the camera and did not look at the monkey. The duration of this stimulus could not be controlled and depended on how long it took the animals to consume the food.

The ***Playback of food calls (PB+)*** consisted of food calls given by one to three members of the subjects social group. Each of the playbacks lasted for 45 seconds. The calls were again recorded with the Avisoft UltraSoundGate 116H laboratory software and a condenser

microphone CM16/CMPA and cut in “iMovie”. To imitate naturally sounding playbacks, the food calls were regularly interrupted by pauses that were similarly arranged for all the individuals. The calls were played back in the same way as the aggressive vocalizations.

Finally, we used a ***control*** stimulus that was assumed to not evoke a strong emotional reaction. The control stimulus was a continuation of the baseline condition with the exception that the experimenter no longer offered a mealworm every 20 sec. Instead, she turned her body to the camera and did not look at the subject. This stimulus lasted for about 45–60 seconds.

The exact durations of the stimuli are listed in table S.2 of the supplementary material.

2.5 Thermal measurement and coding

During the experiments the monkeys were constantly filmed with an infrared thermography camera of the model FLIR T620 (temperature sensitivity: 0.04 °C; resolution: 640 x 480 pixels; sampling rate: 30 fps). For acclimatization, the camera was located in the experimental room and turned on at least 15 minutes before the experiment started. The camera was placed on a tripod in front of the testing compartment about one meter from the mesh. It was controlled with the FLIR Tools+ software from a laptop. The room temperature and humidity were periodically recorded and entered into the camera while the emissivity was constantly set on 0.98. To ensure an optimal and clear image, the orientation of the camera was manually adjusted towards the subject when it was climbing up or down. The focus of the camera was controlled by hand to prevent the autofocus from focusing on the mesh rather than the subject.

For the extraction of the nasal temperature data, we used a customized MATLAB (R2018b) script. To do so we manually placed an ellipse over the region of interest (nose) on the first appropriate frame in each two-second interval and exported the minimal temperature of this area automatically to an excel file. If a respective interval did not contain a useful frame it was omitted. Appropriate frames were selected based on the following criteria (see figure 1 for examples). (1) The subject had to be right behind the frontal mesh to ensure that its face was in the focus of the camera. (2) The subjects’ head had to be oriented straight toward the camera (maximum tilt angle of about 45°). (3) The nasal region was not covered by the mesh. (4) The nose was not blurred due to movements of the subject or the camera. Due to the resolution of the thermography camera, we were not able to recognize the tiny nostrils of the marmosets in the thermographic images. Therefore, we could not exclude potential effects of breathing patterns on individual raw data points. However, we reasoned that a potential change in temperature due to breathing would not influence our main results because it only leads to a change in the amplitude of the oscillation around a mean. This means that inhaling and exhaling compensate each other and hence potential changes in nasal temperature due to the experimental stimuli are still detectable.

20% of all thermography videos were analyzed by a second rater to assess inter-rater reliability. The videos could be coded with very high reliability, resulting in an intraclass correlation coefficient (ICC 3) of 1 with a confidence interval of 1 to 1.

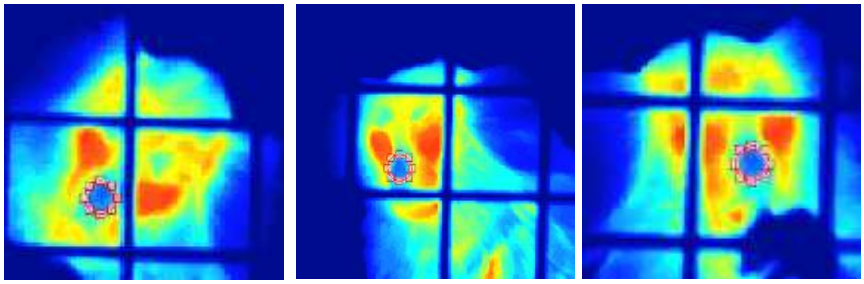


Figure 1 | Example of thermographic frames used to measure nasal temperature. The circles at the nasal region indicate the area from which the minimal temperature was measured

2.6 Behavioral measurement and coding

For the behavioral measurements we additionally filmed the animals with a camcorder (SONY, HDR-CX730E). The camcorder was placed on a tripod right next to the thermography camera. The behavioral analysis was conducted with the software INTERACT (MANGOLD GmbH, version 16).

We assessed piloerection of the tail as a signal of high arousal [43,44]. Therefore, we coded the level of piloerection of the tail every five seconds by assigning one of the three following levels: 0 = no piloerection, 1 = half of the tail shows piloerection or the complete tail shows weak piloerection, 2 = the complete tail shows strong piloerection. If the state of piloerection was not visible at the respective time point, the time interval was extended by two seconds before it was coded as not available (N.A.). Piloerection was coded from 30 seconds before the stimulus onset until the stimulus offset. Finally, to estimate the change in the extent of piloerection during the stimulation phase we subtracted the mean of all scores collected during the baseline phase from the scores of the stimulation phase. For the coding of the piloerection, we reached an intraclass correlation coefficient (ICC 3) of 0.94 with a 95% confidence interval from 0.84 to 0.98 (based on 20% of all the videos coded by two raters).

Finally, we coded the activity level of the monkeys to control for a potential impact of physical activity on nasal temperature. We counted the number of changes in position during the period of 30 seconds before the stimulus onset until 30 seconds after the stimulus had ended. A change in position was defined as the animal moving away all four limbs from the current position and moving at least one body length. Movements on the perch were not counted as changes in position because the perch was quite short (33 cm) and hence physical activity was highly restricted. When a subject changed between two positions via a third position that was occupied for less than ten frames (0.4 s), this was only counted as one change. For example, when a subject was hanging on the ceiling and jumped via the floor directly onto the perch it was counted as one change in position. We calculated a pre (30 seconds before stimulus onset) and a post (stimulus onset until 30 seconds after the stimulus) value by dividing the number of changes in position by the length of the respective period (in minutes) to get frequency measures of activity. For the activity measure, we reached an intraclass correlation coefficient (ICC 3) of 0.96 with a 95% confidence interval from 0.88 to 0.99 (based on 20% of all the videos coded by two raters).

2.7 Data analysis

All statistical analyses were conducted in R 3.4.4. The assumption of homogeneous residuals and the absence of influential outliers were assessed by inspecting residual plots and cook's distances, respectively. Linear mixed models (LMM) were performed with the R function "lme" of the package "nlme". The models were built up in a step-forward procedure, starting with a null model that only included the random effect (random intercept) and adding all the fixed effects step by step (first all the main effects, then the interactions). To assess whether an added factor had a significant impact on the dependent variable, we always compared the respective model to the previous one (the one that did not include the respective factor) by a likelihood ratio test (R function "anova"). If a fixed factor or an interaction did not lead to an improvement of the model it was excluded from the final model if its higher-order interactions were also non-significant [45].

2.7.1 Nasal thermal reaction

To assess temporal changes in nasal temperature, we compared the temperature measures taken during the last 30 sec of the baseline condition (pre) to the measures taken 30 secs after the stimulus onset (post). With this protocol, we would most likely capture the immediate stress response (norepinephrine based "fight or flight" reaction) rather than a slower, potential increase in cortisol [20,21].

Since we were interested in temperature changes relative to baseline, rather than in absolute temperature, we calculated a centered temperature measure, by subtracting the individual mean temperature of the baseline from all temperature measures of the respective session.

To examine the effect of the different stimuli on changes in nasal temperature we conducted a LMM (model 1) on the centered nasal temperature. We set condition (food, teasing, PB-, PB+ and control), phase (pre and post), sex (male, female) as well as all two- and three-way interactions as fixed effects. Further, we added the subject as random effect to control for the repeated measures within a session.

To assess in more detail how males and females reacted to the different stimuli, we compared least square means of nasal temperature between pre and post phase per condition and sex based on the final model. Finally, to investigate the thermal reactions on the individual level, we performed Mann-Whitney U tests that compared nasal temperature between the pre and post-phase separately within each session (see supplementary S.3, table S.3).

2.7.2 Piloerection

First, we assessed the effect of the different stimuli on piloerection of the tail by calculating a LMM (model 2) on the extent of piloerection of the tail that included condition, sex as well as their interaction as fixed effect and subject as random effect. Based on the final model we compared least square means among the conditions.

To examine the link between the nasal thermal reaction and the piloerection, we set up a LMM (model 3) on the temperature data of the post-phase. We included the extent of piloerection of the tail, sex and their interaction as fixed effect and subject as random effect.

2.7.3 Physical activity

First, we assessed whether the stimuli had an effect on physical activity. Therefore, we ran a LMM (model 3) on physical activity (frequency of changes in position) that included condition, phase, sex as well as all their interactions as fixed effects and the subject as random effect. Based on the final model, we compared least square means between the phases within each condition.

In a second step we examined whether the observed temperature changes could be mere artefacts of changes in body temperature due to physical activity. First, we examined whether changes in temperature were better explained by experimental condition, physical activity, or piloerection of the tail. We therefore compared the AIC values of three LMMs that we conducted on nasal temperature of the post phase (the data of one male marmoset (Merkur) in the food call condition was excluded as no piloerection data was available) that each included one of the three fixed factors of interest. Subject was always set as random effect.

Second, we looked for each experimental condition separately if there was a link between physical activity and nasal temperature. We thus ran five separate LMMs, one for each condition. We used nasal temperature of the post-phase as dependent variable while physical activity (post value), sex as well as their interaction were set as fixed effects and the subject as random effect.

3 Results

3.1 Nasal thermal reaction

The best model to explain nasal temperature included the factors condition, phase, sex as well as all their two- and three-way interactions (see table 2, model 1). We used pairwise comparisons of least square means of nasal temperature between the pre and post phase per condition and sex to disentangle these effects (table 3 and figure 2). In the negative conditions (teasing and playback of aggressive vocalizations), both males and females showed a strong decrease in nasal temperature in response to the stimulus. The temperature drops were more pronounced in females than males. In the positive conditions (preferred food and playback of food calls), males and females reacted differently. Males showed an increase in both conditions, whereas females showed no change in temperature after eating a preferred food item and a decrease in temperature after the playback of food calls. In the control condition, finally, females showed no significant change in temperature whereas males showed a modest increase. See supplementary S.3 for an overview over the results at the individual level, separately for each condition.

Table 3 | Pairwise comparisons of least square means of nasal temperature between pre and post per condition and sex. Bold *p*-values indicate a significant result (< 0.05).

Final model: Nasal temperature ~ Condition*Phase*Sex + Subject (random effect)							
Contrast	**Condition**	**Sex**	**Estimate (°C)**	**SE (°C)**	**df**	**t-value**	***p*-value**
Pre to Post phase	Teasing	Male	- 0.524	0.082	1012	-6.361	**< 0.0001**
		Female	- 0.832	0.068	1012	-12.188	**< 0.0001**
	PB-	Male	- 0.215	0.073	1012	-2.936	**0.003**
		Female	- 0.926	0.073	1012	-12.703	**< 0.0001**
	Food	Male	0.202	0.076	1012	2.650	**0.008**
		Female	0.028	0.083	1012	0.342	0.733
	PB+	Male	0.244	0.083	1012	2.935	**0.003**
		Female	- 0.238	0.084	1012	-2.820	**0.005**
	Control	Male	0.174	0.079	1012	2.196	**0.028**
		Female	- 0.143	0.082	1012	-1.744	0.081

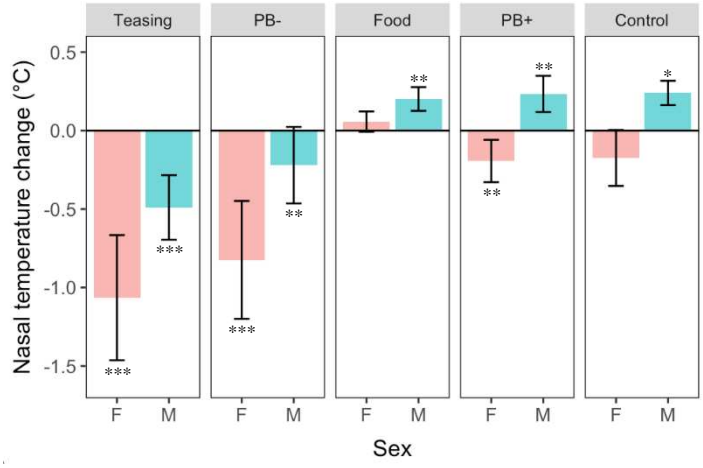


Figure 2 | Grand mean of nasal temperature changes (temperatures of post phase) during the five different stimuli separately for each sex (error bars indicate SE). Significant changes are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

3.2 Piloerection

First, we examined how the extent of piloerection differed between the conditions and the sexes. The respective LMM (see table 2, model 2) showed a significant effect of condition on the extent of piloerection, whereas sex and the interaction of sex and condition were not significant. Based on the final model, pairwise multiple comparisons of least square means of piloerection among the conditions (see table 4 and figure 3) revealed that in the negative conditions, the teasing and the playback of aggressive vocalizations, the monkeys showed significantly more piloerection than in the other conditions. Between the two negative conditions the extent of piloerection was not significantly different.

Table 4 | Multiple pairwise comparisons of least square means of the extent of piloerection among the conditions. Bold *p*-values indicate a significant differences (< 0.05)

Final model: Extent of piloerection ~ Condition + Subject (random effect)						
Contrast		**Estimate**	**SE**	**df**	**t-value**	***p*-value***
Teasing	PB-	-0.023	0.154	43	-0.150	1.000
	Food	0.619	0.159	43	3.883	**0.003**
	PB+	0.544	0.171	43	3.183	**0.022**
	Control	0.658	0.163	43	4.038	**0.002**
PB-	Food	0.642	0.154	43	4.168	**0.001**
	PB+	0.567	0.166	43	3.417	**0.012**
	Control	0.681	0.158	43	4.320	**0.001**
Food	PB+	-0.075	0.171	43	-0.436	0.992
	Control	0.039	0.163	43	0.237	0.999
PB+	Control	0.113	0.174	43	0.649	0.966

* *p*-values adjusted by tukey method

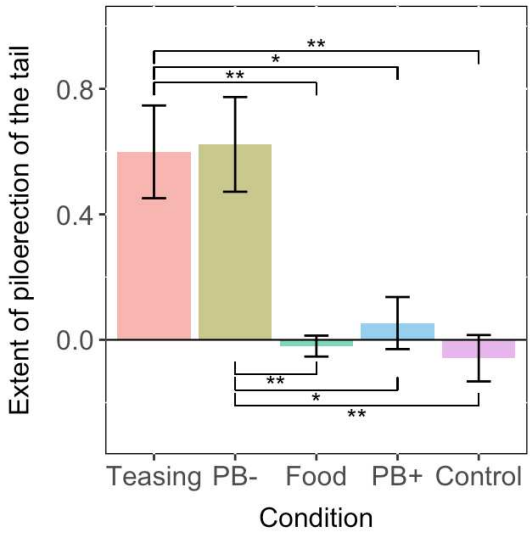


Figure 3 | Mean of the extent of piloerection of the tail for all the different conditions. Significant differences between the conditions are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

To validate the findings of the thermoimaging data, we tested whether there was a link between the extent of piloerection and the change in nasal temperature. A LMM (table 2, model 3) conducted on nasal temperature revealed that the extent of piloerection is a significant predictor of the change in nasal temperature: more piloerection of the tail was linked to a stronger decrease in nasal temperature (see figure 4). Although sex did significantly improve the model, which was expected given model 1, the interaction of the extent of piloerection and sex was not a significant predictor of the thermal reaction.

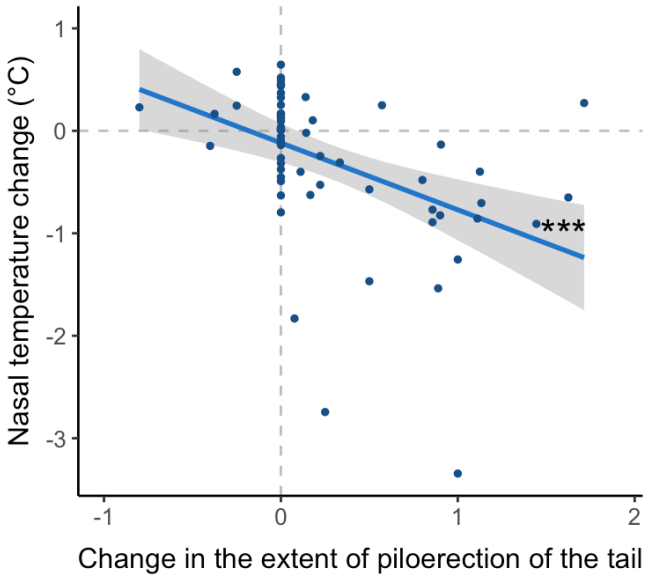


Figure 4 | Relationship between piloerection of the tail and nasal thermal reaction. Each point indicates the mean data of one session (one subject in one condition). X-values less than zero indicate a decrease while positive values indicate an increase in piloerection of the tail from the baseline to the stimulation phase. The grey band represents the standard error and the significance is indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

3.3 Physical activity

In a first analysis we examined whether the experimental stimuli had an influence on physical activity. We found that condition, phase and their interaction had an effect, whereas sex as well as all its interactions were no significant predictors for physical activity (table 2, model 4). Based on the final model, we compared least square means of nasal temperature between the pre and post phase within each condition (table 5, figure 5) and found a significant increase in physical activity during the playback of aggressive vocalizations, but no change in any of the other conditions.

Table 5 | Comparisons of least square means of physical activity between pre and post within each condition. Bold *p*-values indicate a significant result (< 0.05).

Final model: Physical activity ~ Condition*Phase + Subject (random effect)						
Contrast	**Condition**	**Estimate**	**SE**	**df**	**t-value**	***p*-value**
Pre to Post phase	Teasing	1.369	0.990	112	1.385	0.169
	PB-	**4.606**	0.925	112	4.980	**<.0001**
	Food	-0.481	0.990	112	- 0.487	0.627
	PB+	1.264	1.068	112	1.183	0.239
	Control	1.642	1.026	112	1.601	0.112

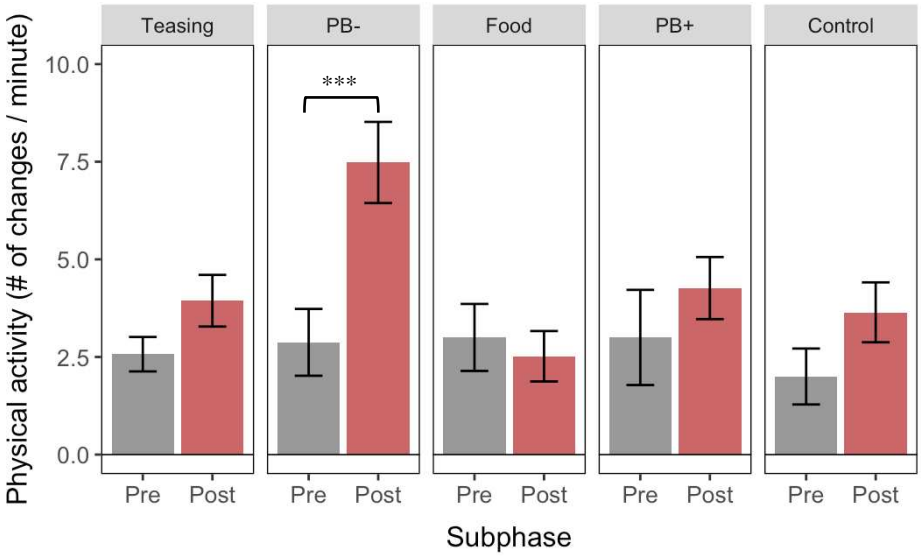


Figure 5 | Mean of physical activity for the pre and post phase for each condition (error bars indicate SE). Significant changes are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

Second, we were interested in the impact of physical activity on nasal temperature. We thus analyzed whether the change in nasal temperature was better explained by mere physical activity rather than the emotional state. Based on the findings above, emotional state was proxied by experimental condition and piloerection of the tail. The comparison of the AIC values of the three LMMs revealed that the thermal reaction was explained worst by physical activity, better by piloerection, and best by condition (table 6).

Table 6 | Comparison of the relative impact of the fixed factors condition, extent of piloerection and physical activity on nasal temperature by AIC values.

Dependent variable	**Fixed factor**	**ΔAIC**
Nasal temperature (Post phase, all conditions)	Condition	0
	Extent of piloerection	76.265
	Activity (post value)	135.3

Finally, we quantified the effect of physical activity on nasal temperature for each condition separately with LMMs (table 2, model 5 a–e; figure 6). There was no consistent relationship between physical activity and nasal temperature among the different conditions, except in the aggressive playback condition, where higher physical activity was significantly correlated with a stronger decrease in nasal temperature.

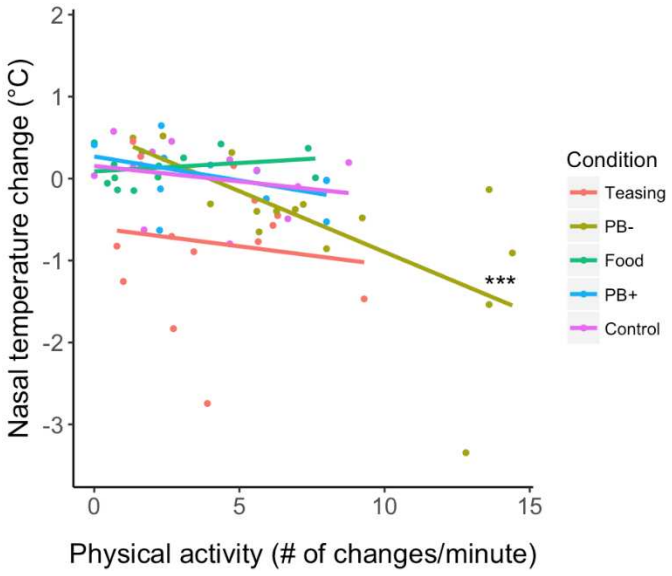


Figure 6 | Relationship between physical activity (post value) and nasal thermal reaction in each condition. Each point indicates the mean data of one session (one subject in one condition). Significant relationships are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

506

507

Table 2 | Summary of all linear mixed models. The df, χ^2 and p-values resulted from the respective model comparisons by likelihood ratio tests. The order in which the different fixed factors were added to the model corresponds to the order in which the factors are listed here. In all the models, subject was set as random factor.
Bold p-values indicate significant results (< 0.05)

Model		Dependent variable	Fixed factor / effect	df	χ^2	*p*-value
1		Nasal temperature (Pre and Post phase; all conditions)	Condition	4	119.907	**< 0.0001**
			Phase	1	83.200	**< 0.0001**
			Sex	1	4.983	**0.026**
			Condition*Phase	4	172.789	**<.0001**
			Condition*Sex	4	2.348	0.672
			Phase*Sex	1	65.729	**<0.0001**
			Condition*Phase*Sex	4	15.134	**0.004**
2		Extent of piloerection (all conditions)	Condition	4	30.888	**<0.0001**
			Sex	1	1.433	0.231
			Condition*Sex	4	2.621	0.623
3		Nasal temperature (Post phase, all conditions)	Extent of piloerection	1	110.794	**<0.0001**
			Sex	1	4.927	**0.026**
			Extent of piloerection*Sex	1	0.551	0.458
4		Activity (Pre and Post phase; all conditions)	Condition	4	12.447	**0.014**
			Phase	1	14.400	**0.000**
			Sex	1	0.669	0.413
			Condition*Phase	4	15.285	**0.004**
			Condition*Sex	4	2.756	0.600
			Phase*Sex	1	1.288	0.256
			Condition*Phase*Sex	4	1.599	0.809
5	a) Teasing	Nasal temperature (Post phase; Teasing)	Physical activity (post value)	1	0.248	0.619
	b) PB-	Nasal temperature (Pre and post phase; PB-)		1	9.572	**0.002**
	c) Food	Nasal temperature (Pre and post phase; Food)		1	0.769	0.380
	d) PB+	Nasal temperature (Pre and post phase; PB+)		1	2.494	0.114
	e) Control	Nasal temperature (Pre and post phase; Control)		1	0.832	0.362

4 Discussion

Our goal was to evaluate the usefulness of infrared thermography as a non-invasive and objective measure of emotional reactions in common marmosets. More precisely, we focused on how nasal temperature changed in response to different stimuli that were chosen to induce varying types of emotional reactions. As predicted, we found that nasal temperature dropped significantly in situations that were assumed to be highly arousing. To validate whether the drop in nasal temperature indeed reflected arousal, we first analyzed whether it was correlated with an independent behavioral measure of arousal in marmosets, namely piloerection of the tail. We found such a correlation, which was a first indication that a decrease in nasal temperature indeed reflects arousal. However, an alternative possibility is that the decrease in nasal temperature is the result of an increase in physical activity, rather than a marker of arousal per se. To exclude this alternative, we also measured the individuals' activity. Our results show that there was no consistent relationship between physical activity and the thermal response across conditions. In fact, we only found a significant correlation between these two measures in the aggressive playback condition. Even though in principle, activity in this condition may be responsible for the decrease in nasal temperature, the lack of such a correlation in all other conditions suggest that the temperature change cannot be solely explained by activity. Most likely, the negative stimulus in the aggressive playback condition led to an emotional reaction reflected in nasal temperature changes and a behavioral reaction (increased activity). The latter may also have influenced nasal temperature and therefore, the observed thermal signature may indicate the result of both internal and behavioral (activity) consequences of arousal. Finally, there was no increase in physical activity in the teasing condition, where we observed the strongest decrease in nasal temperature, corroborating that most variation in nasal temperature is independent of activity.

Accordingly, changes in nasal temperature were far better explained by the stimulus type (condition) or by piloerection rather than by activity. It is thus unlikely that changes in nasal temperature were a mere reflection of physical activity, at least in the current setting. Admittedly, in more naturalistic settings with freely moving individuals, such as in the wild, variation in activity will be much greater and may well have a stronger influence on nasal temperature. Nevertheless, these results support that infrared thermography is a reliable and valid measure of arousal in captive common marmosets that are tested under controlled conditions.

The negative stimuli, i.e. teasing and a playback of aggressive vocalizations, elicited the most obvious thermal responses. These two stimuli were a priori assumed to be perceived as negative in valance and highly arousing. In both conditions nasal temperature decreased significantly after the stimulus. This result was true for most of the individuals and although females reacted more strongly, the effect was present in both sexes. Furthermore, the amplitude of the thermal responses was relatively large compared to the ones in the other conditions. In addition to the thermal reaction, piloerection of the tail was significantly more prevalent in the two negative conditions compared to the others. This indicates that the teasing and the playback of aggressive vocalizations were indeed more arousing than the other conditions. Furthermore, there was also a negative correlation between the extent of piloerection of the tail and nasal temperature. Overall, these observations are consistent with the pattern found in other non-

564 human primates, namely that high negative arousal is accompanied by a decrease in nasal
565 temperature [3,6,8,11,19].
566 In contrast to the negative conditions, the temperature responses to the two positive and
567 the control stimulus were less pronounced, suggesting that these stimuli had a smaller impact
568 on the emotional state. The piloerection data was fully consistent with this idea: while
569 piloerection was abundant in the negative conditions, it was almost absent in the positive and
570 the control conditions, without any sex differences. The behavioral data would thus suggest no
571 emotional reaction in the positive and the control conditions, but the thermal data revealed a
572 more fine-grained pattern: an increase in nasal temperature was only observed in the males,
573 whereas there was no effect (preferred food and control condition) or even a temperature
574 decrease (food calls condition) in the females. The patterns observed in the positive conditions
575 as well as the slight sex differences in the negative conditions (females showed a stronger
576 temperature decrease) go in line with the predicted link between arousal and nasal temperature
577 if we consider behavioral sex differences in marmosets. In general, female common marmosets
578 seem to be more interested and motivated than males in situations where food is involved. They
579 are faster and more persistent when they have to search for food [46] or when they have to solve
580 a food-related task [43,47] and they are more likely to aggressively defend their food against
581 group members [48]. Further females show higher intra-sexual competition and are more
582 involved in between-group aggression [49,50].
583 The lack of an increase in nasal temperature following the food stimulus may thus indicate that
584 the females were slightly aroused by the anticipation of food competition, which overrode the
585 relaxing effect of the food per se. In addition, the temperature decrease in the food call condition
586 in females may simply reflect the excitement about the presence of high-quality food as
587 announced by the food calls. The higher food motivation may also explain the stronger decrease
588 in temperature in females during the teasing. On the other hand, the stronger reaction of the
589 females in the aggressive playback condition may be caused by the fact that they were more
590 aroused due to potential encounters with same-sex competitors.
591 In the males, there appeared to be a consistent overall pattern in the two positive and the control
592 conditions: males tended to have higher nasal temperatures after all these stimuli. This could
593 suggest that the males were more likely to be not fully habituated to the situation by the start of
594 the experiment and would have needed more time to calm down at the start of an experimental
595 session. This is not unlikely since among callitrichids, males are known to be the more vigilant
596 sex and they are more responsible for the detection of potential dangers to their group [51,52],
597 what could explain a higher level of arousal and hence a slower habituation during the
598 separation procedure. However, additional analyses suggest that this explanation doesn't hold,
599 because males did not show a steeper positive increase in nasal temperature during the baseline
600 than females (see supplementary materials S.4 for more details). Therefore, a more plausible
601 explanation would be that males perceive these three stimuli as relaxing. This was expected for
602 the food and the food calls playback condition but not for the control condition. However,
603 during the control condition, the baseline phase where the experimenter interacted with the
604 subject by giving him a mealworm every 20 seconds was followed by the stimulus phase where
605 the experimenter directed her attention away from the subject and instead watched at the
606 camera. It is possible that this removal of attention from the experimenter reduced tension and
607 thus arousal in the males. In fact, this would be consistent with a previous study that showed
608 that male marmosets are more emotionally aroused than females when confronted with a human

experimenter [43]. Whether the increases in nasal temperature in the two positive conditions were indeed evoked by the stimuli itself or were again caused by the distance to the experimenter is not clear. However, both possibilities support that a reduction of the arousal level is linked to an increase in nasal temperature.

Despite the clear sex differences in the overall data, some individual females showed the male pattern of thermal changes, and some individual males showed the female pattern (see supplementary material S.3). This suggests that in addition to systematic sex differences, individual differences exist with regard to how a specific situation is perceived. Importantly, these could only be made visible with thermo-imaging but not with measuring piloerection of the tail, a traditionally used behavioral indicator of arousal. Thermography thus appears to be a particularly sensitive measure of arousal in marmosets. High sensitivity of thermography to quantify arousal had previously been proposed for humans by Kosonogov et al. [22]. They compared nasal temperature changes with the skin conductance response and also found that the thermal reaction was more sensitive to arousal than the other measure. Nevertheless, since interindividual differences in physiological and behavioral responses to emotions are possible, it is most desirable to code several indicators of emotional arousal simultaneously.

We only focused on the nasal region due to its clear visibility in the thermal images, which made the tracking of the respective temperature highly reliable. However, in future studies, the thermal reactions of additional body parts could be investigated. Results in humans and other primates indeed suggest that combinations of different thermal reactions can distinguish between different emotions in more detail [3,13,33,53]. Next to measuring peripheral body regions, one promising area are the eyes, that have been shown to change their temperature during emotional reactions in humans and different animals [20,54–57].

Overall, the current experiment supports that in common marmosets nasal temperature is sensitive to emotional changes, and that infrared thermography is an appropriate, highly objective and non-invasive method to infer these thermal reactions. In fact, this method appears to be more sensitive than the traditional behavioral marker of arousal, i.e. piloerection. Furthermore, our continuous graphical representations of marmoset thermal reactions (see supplementary material S.3, figure S.1) provide useful benchmarks about onset latencies and the amplitude of the thermal reaction for future studies.

Mirroring the literature, particularly consistent results were achieved with highly arousing, negative stimuli, most likely because they are perceived in the same way by all individuals. Equally universal positive stimuli are more difficult to identify, and even the presentation of highly preferred food may result in negative arousal if individuals anticipate food competition. The measurement of valence independently of arousal thus continues to be a challenge.

Acknowledgments

We thank Erik Willems for statistical support, Gisep Bazzell for assistance and animal care taking, and Carel van Schaik for providing useful inputs on the manuscript. This project was supported by a SNF grant to JMB (grant number 31003A_172979) and the Janggen Poehn Stiftung (RKB).

Data accessibility

Data can be accessed from the Mendeley Data repository [58].

Competing interests

We declare we have no competing interests.

References

[1] E.S. Paul, E.J. Harding, M. Mendl, Measuring emotional processes in animals: The utility of a cognitive approach, Neurosci. Biobehav. Rev. 29 (2005) 469–491. doi:10.1016/j.neubiorev.2005.01.002.

[2] M. Mendl, O.H.P. Burman, R.M.A. Parker, E.S. Paul, Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms, Appl. Anim. Behav. Sci. 118 (2009) 161–181. doi:10.1016/j.applanim.2009.02.023.

[3] G. Dezecache, K. Zuberbühler, M. Davila-Ross, C.D. Dahl, Skin temperature changes in wild chimpanzees upon hearing vocalizations of conspecifics, R. Soc. Open Sci. 4 (2017) 160816. doi:10.1098/rsos.160816.

[4] J. Clay-Warner, D.T. Robinson, Infrared thermography as a measure of emotion response, Emot. Rev. 7 (2015) 157–162. doi:10.1177/1754073914554783.

[5] S.T. Boysen, G.G. Berntson, Conspecific recognition in the chimpanzee (Pan troglodytes): cardiac responses to significant others, J. Comp. Psychol. 103 (1989) 215–220. doi:10.1037/0735-7036.103.3.215.

[6] F. Kano, S. Hirata, T. Deschner, V. Behringer, J. Call, Nasal temperature drop in response to a playback of conspecific fights in chimpanzees: A thermo-imaging study, Physiol. Behav. 155 (2016) 83–94. doi:10.1016/j.physbeh.2015.11.029.

[7] M. Katayama, T. Kubo, K. Mogi, K. Ikeda, M. Nagasawa, T. Kikusui, Heart rate variability predicts the emotional state in dogs, Behav. Processes. 128 (2016) 108–112. doi:10.1016/j.beproc.2016.04.015.

[8] K. Kuraoka, K. Nakamura, The use of nasal skin temperature measurements in studying emotion in macaque monkeys, Physiol. Behav. 102 (2011) 347–355. doi:10.1016/j.physbeh.2010.11.029.

[9] E.A. Butler, T.L. Lee, J.J. Gross, Does expressing your emotions raise or lower your blood pressure?: The answer depends on cultural context, J. Cross. Cult. Psychol. 40 (2009) 510–517. doi:10.1177/0022022109332845.

[10] J.M. Wood, C.R. Schnell, F. Cumin, J. Menard, R.L. Webb, Aliskiren, a novel, orally effective renin inhibitor, lowers blood pressure in marmosets and spontaneously hypertensive rats., J. Hypertens. 23 (2005) 417–26. doi:10.1097/00004872-200502000-00025.

[11] S. Ioannou, H. Chotard, M. Davila-Ross, No strings attached: physiological monitoring of rhesus monkeys (Macaca mulatta) with thermal imaging, Front. Behav. Neurosci. 9 (2015) 1–10. doi:10.3389/fnbeh.2015.00160.

[12] C.T. Miller, W.A. Freiwald, D.A. Leopold, J.F. Mitchell, A.C. Silva, X. Wang, Marmosets: A Neuroscientific Model of Human Social Behavior, Neuron. 90 (2016) 219–233. doi:10.1016/j.neuron.2016.03.018.

[13] S. Ioannou, V. Gallese, A. Merla, Thermal infrared imaging in psychophysiology: Potentialities and limits, Psychophysiology. 51 (2014) 951–963. doi:10.1111/psyp.12243.

[14] S. Ioannou, S. Ebisch, T. Aureli, D. Bafunno, H.A. Ioannides, D. Cardone, B. Manini, G.L. Romani, V. Gallese, A. Merla, The autonomic signature of guilt in children: A thermal infrared imaging study, PLoS One. 8 (2013) e79440. doi:10.1371/journal.pone.0079440.

[15] D.M.L. Vianna, P. Carrive, Changes in cutaneous and body temperature during and after conditioned fear to context in the rat, Eur. J. Neurosci. 21 (2005) 2505–2512. doi:10.1111/j.1460-9568.2005.04073.x.

[16] A. Kistler, C. Mariauzouls, K. Von Berlepsch, Fingertip temperature as an indicator for sympathetic responses, Int. J. Psychophysiol. 29 (1998) 35–41. doi:10.1016/S0167-8760(97)00087-1.

[17] C.K. Or, V.G. Duffy, Development of a facial skin temperature-based methodology for

non-intrusive mental workload measurement, Occup. Ergon. 7 (2007) 83–94. http://www.scopus.com/inward/record.url?eid=2-s2.0-34547668616&partnerID=tZOtx3y1.
[18] S.J. Ebisch, T. Aureli, D. Bafunno, D. Cardone, G.L. Romani, A. Merla, Mother and child in synchrony: Thermal facial imprints of autonomic contagion, Biol. Psychol. 89 (2012) 123–129. doi:10.1016/j.biopsycho.2011.09.018.
[19] K. Nakayama, S. Goto, K. Kuraoka, K. Nakamura, Decrease in nasal temperature of rhesus monkeys (Macaca mulatta) in negative emotional state, Physiol. Behav. 84 (2005) 783–790. doi:10.1016/j.physbeh.2005.03.009.
[20] M. Stewart, G.A. Verkerk, K.J. Stafford, A.L. Schaefer, J.R. Webster, Noninvasive assessment of autonomic activity for evaluation of pain in calves, using surgical castration as a model, J. Dairy Sci. 93 (2010) 3602–3609. doi:10.3168/jds.2010-3114.
[21] B.S. McEwen, R.M. Sapolsky, Stress and cognitive function, Curr. Opin. Neurobiol. 5 (1995) 205–216. doi:10.1016/0959-4388(95)80028-X.
[22] V. Kosonogov, L. De Zorzi, J. Honoré, E.S. Martínez-Velázquez, J.L. Nandrino, J.M. Martinez-Selva, H. Sequeira, Facial thermal variations: A new marker of emotional arousal, PLoS One. 12 (2017) 1–15. doi:10.1371/journal.pone.0183592.
[23] E. Salazar-López, E. Domínguez, V. Juárez Ramos, J. de la Fuente, A. Meins, O. Iborra, G. Gálvez, M.A. Rodríguez-Artacho, E. Gómez-Milán, The mental and subjective skin: Emotion, empathy, feelings and thermography, Conscious. Cogn. 34 (2015) 149–162. doi:10.1016/j.concog.2015.04.003.
[24] R. Nakanishi, K. Imai-Matsumura, Facial skin temperature decreases in infants with joyful expression, Infant Behav. Dev. 31 (2008) 137–144. doi:10.1016/j.infbeh.2007.09.001.
[25] P.A. Boudewyns, A comparison of the effects of stress vs. relaxation instruction on the finger temperature response, Behav. Ther. 7 (1976) 54–67. doi:10.1016/S0005-7894(76)80219-5.
[26] S.E. Rimm-Kaufman, J. Kagan, The psychological significance of changes in skin temperature, Motiv. Emot. 20 (1996) 63–78. doi:10.1007/BF02251007.
[27] A.C. Hahn, R.D. Whitehead, M. Albrecht, C.E. Lefevre, D.I. Perrett, Hot or not? Thermal reactions to social contact, Biol. Lett. 8 (2012) 864–867. doi:10.1098/rsbl.2012.0338.
[28] A. Merla, G.L. Romani, Thermal signatures of emotional arousal: A functional infrared imaging study, in: Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc., 2007: pp. 247–249. doi:10.1109/IEMBS.2007.4352270.
[29] B.N. Lang, P.J., Bradley, M.M., & Cuthbert, International affective picture system (IAPS): Affective ratings of pictures and instruction manual., Tech. Rep. A-8. (2008). doi:10.1016/j.epsr.2006.03.016.
[30] N. Ludwig, M. Gargano, F. Luzi, C. Carenzi, M. Verga, Technical note: Applicability of infrared thermography as a non invasive measurement of stress in rabbit, World Rabbit Sci. 15 (2007) 199–205. doi:10.4995/wrs.2007.588.
[31] S. Riemer, L. Assis, T.W. Pike, D.S. Mills, Dynamic changes in ear temperature in relation to separation distress in dogs, Physiol. Behav. 167 (2016) 86–91. doi:10.1016/j.physbeh.2016.09.002.
[32] L.C. Grandi, E. Heinzl, Data on thermal infrared imaging in laboratory non-human primates: Pleasant touch determines an increase in nasal skin temperature without affecting that of the eye lachrymal sites, Data Br. 9 (2016) 536–539. doi:10.1016/j.dib.2016.09.029.
[33] H. Chotard, S. Ioannou, M. Davila-Ross, Infrared thermal imaging: Positive and negative emotions modify the skin temperatures of monkey and ape faces, Am. J. Primatol. 80 (2018) e22863. doi:10.1002/ajp.22863.

765 [34] J. Walker, J. MacLean, N.G. Hatsopoulos, The marmoset as a model system for
766 studying voluntary motor control, Dev. Neurobiol. 77 (2017) 273–285.
767 doi:10.1002/dneu.22461.
768 [35] J.F. Mitchell, D.A. Leopold, The marmoset monkey as a model for visual
769 neuroscience, Neurosci. Res. 93 (2015) 20–46. doi:10.1016/j.neures.2015.01.008.
770 [36] N. Kishi, K. Sato, E. Sasaki, H. Okano, Common marmoset as a new model animal for
771 neuroscience research and genome editing technology, Dev. Growth Differ. 56 (2014)
772 53–62. doi:10.1111/dgd.12109.
773 [37] J.M. Burkart, E. Guerreiro Martins, F. Miss, Y. Zürcher, From sharing food to sharing
774 information, Interact. Stud. Stud. Soc. Behav. Commun. Biol. Artif. Syst. 19 (2018)
775 136–150. doi:10.1075/is.17026.bur.
776 [38] J.M. Burkart, C. Finkenwirth, Marmosets as model species in neuroscience and
777 evolutionary anthropology, Neurosci. Res. 93 (2015) 8–19.
778 doi:10.1016/j.neures.2014.09.003.
779 [39] R.K. Brügger, T. Kappeler-Schmalzriedt, J.M. Burkart, Reverse audience effects on
780 helping in cooperatively breeding marmoset monkeys, Biol. Lett. 14 (2018) 20180030.
781 doi:10.1098/rsbl.2018.0030.
782 [40] M. Rukstalis, J.A. French, Vocal buffering of the stress response: Exposure to
783 conspecific vocalizations moderates urinary cortisol excretion in isolated marmosets,
784 Horm. Behav. 47 (2005) 1–7. doi:10.1016/j.yhbeh.2004.09.004.
785 [41] A. Vitale, M. Zanzoni, A. Queyras, F. Chiarotti, Degree of social contact affects the
786 emission of food calls in the common marmoset (Callithrix jacchus), Am. J. Primatol.
787 59 (2003) 21–28. doi:10.1002/ajp.10060.
788 [42] Y. Zürcher, J.M. Burkart, Evidence for Dialects in Three Captive Populations of
789 Common Marmosets (Callithrix jacchus), Int. J. Primatol. 38 (2017) 780–793.
790 doi:10.1007/s10764-017-9979-4.
791 [43] M.N. Schubiger, F.L. Wüstholz, A. Wunder, J.M. Burkart, High emotional reactivity
792 toward an experimenter affects participation, but not performance, in cognitive tests
793 with common marmosets (Callithrix jacchus), Anim. Cogn. 18 (2015) 701–712.
794 doi:10.1007/s10071-015-0837-5.
795 [44] A. Dettling, C.R. Pryce, R.D. Martin, M. Döbeli, Physiological responses to parental
796 separation and a strange situation are related to parental care received in juvenile
797 Goeldi's monkeys (Callimico goeldii), Dev. Psychobiol. 33 (1998) 21–31.
798 doi:10.1002/(SICI)1098-2302(199807)33:1<21::AID-DEV3>3.0.CO;2-U.
799 [45] A. Unwin, Discovering Statistics Using R by Andy Field, Jeremy Miles, Zoë Field, Int.
800 Stat. Rev. 81 (2013) 169–170. doi:10.1111/insr.12011_21.
801 [46] A.M. Michels, Sex differences in food acquisition and aggression in captive common
802 marmosets (Callithrix jacchus), Primates. 39 (1998) 549–556.
803 doi:10.1007/BF02557576.
804 [47] M.E. Yamamoto, C. Domeniconi, H. Box, Sex differences in common Marmosets
805 (Callithrix jacchus) in response to an unfamiliar food task, Primates. 45 (2004) 249–
806 254. doi:10.1007/s10329-004-0088-6.
807 [48] S.D. Tardif, C.B. Richter, Competition for a desired food in family groups of the
808 common marmoset (Callithrix jacchus) and the cotton-top tamarin (Saguinus oedipus).,
809 Lab. Anim. Sci. 31 (1981) 52–5. http://www.ncbi.nlm.nih.gov/pubmed/6789003.
810 [49] C. Lazaro-Perea, Intergroup interactions in wild common marmosets, Callithrix
811 jacchus: Territorial defence and assessment of neighbours, Anim. Behav. 62 (2001)
812 11–21. doi:10.1006/anbe.2000.1726.
813 [50] G. Epple, Maintenance, breeding, and development of marmoset monkeys
814 (Callithricidae) in captivity, Folia Primatol. 12 (1970) 56–76. doi:10.1159/000155281.
815 [51] A. Koenig, Visual scanning by common marmosets (Callithrix jacchus): Functional

aspects and the special role of adult males, Primates. 39 (1998) 85–90. doi:10.1007/BF02557746.
[52] A. Savage, C.T. Snowdon, L.H. Giraldo, L.H. Soto, Parental Care Patterns and Vigilance in Wild Cotton-Top Tamarins (Saguinus oedipus), Adapt. Radiations Neotrop. Primates. (2011) 187–199. doi:10.1007/978-1-4419-8770-9_11.
[53] I.A. Cruz-Albarran, J.P. Benitez-Rangel, R.A. Osornio-Rios, L.A. Morales-Hernandez, Human emotions detection based on a smart-thermal system of thermographic images, Infrared Phys. Technol. 81 (2017) 250–261. doi:10.1016/j.infrared.2017.01.002.
[54] I. Pavlidis, J. Levine, P. Baukol, Thermal image analysis for anxiety detection, in: 2002: pp. 315–318. doi:10.1109/icip.2001.958491.
[55] T. Travain, E.S. Colombo, E. Heinzl, D. Bellucci, E. Prato Previde, P. Valsecchi, Hot dogs: Thermography in the assessment of stress in dogs (Canis familiaris)-A pilot study, J. Vet. Behav. Clin. Appl. Res. 10 (2015) 17–23. doi:10.1016/j.jveb.2014.11.003.
[56] M. Stewart, J.R. Webster, G.A. Verkerk, A.L. Schaefer, J.J. Colyn, K.J. Stafford, Non-invasive measurement of stress in dairy cows using infrared thermography, Physiol. Behav. 92 (2007) 520–525. doi:10.1016/j.physbeh.2007.04.034.
[57] E. Bartolomé, M.J. Sánchez, A. Molina, A.L. Schaefer, I. Cervantes, M. Valera, Using eye temperature and heart rate for stress assessment in young horses competing in jumping competitions and its possible influence on sport performance, Animal. 7 (2013) 2044–2053. doi:10.1017/S1751731113001626.
[dataset] [58] F.A. Ermatinger, R.K. Brügger, J.M. Burkart, Data for: The use of infrared thermography to investigate emotions in common marmosets, Mendeley Data, 2019, (xxx (global persistent identifier))

S.1 List of subjects

Table S.1: Families of the tested marmosets with information about sex, status and age. The subjects that were used in the experiments are marked in black, while grey and cursive indicate family members that were excluded due to their age or insufficient habituation to the experimental set-up.

Group	Subject	Sex	Status	Birth date
1	Mina	f	breeder	10.07.06
	John	m	breeder	10.03.07
	Merkur	m	helper	11.07.12
	Manuka	f	helper	04.05.15
	Mojita	*f*	*immature*	*04.05.15*
2	*Marvin*	*m*	*breeder*	*03.10.06*
	Jet	m	helper	03.04.09
	Jupie	f	helper	29.08.09
	Joyce	f	helper	29.08.09
	James	*m*	*helper*	*21.04.13*
	Jamaica	*f*	*immature*	*10.11.15*
	Jambi	*m*	*immature*	*10.11.15*
3	Tabor	m	-	30.10.08
	Thilo	m	-	06.03.08
4	Jaja	f	breeder	29.08.09
	Membo	*m*	*breeder*	*17.06.09*
	Jandira	m	helper	27.01.15
	Jala	f	helper	27.01.15
	Jelly	*f*	*helper*	*01.07.16*
	Jam	m	helper	01.07.16
5	Lancia	f	breeder	26.08.02
	Lexus	m	breeder	19.07.03
	Lotus	*m*	*helper*	*03.07.12*
	Lola	*f*	*helper*	*29.03.13*
6	Vesta	f	-	05.10.04
	Vito	*m*	-	*30.05.06*
7	*Lea*	*f*	*breeder*	*18.08.07*
	Kyros	m	breeder	03.01.08
	Lima	*f*	*helper*	*15.02.15*

S.2 Durations of the stimuli

Table S.2: Lengths of the stimuli used for all the sessions

Stimulus	Subject	Duration (s)
Control	Jaja	60
	Jala	48
	Jet	59
	Joyce	46
	Jupie	49
	Kyros	60
	Lancia	59
	Lexus	58
	Merkur	60
	Mina	45
	Tabor	45
	Thilo	45
	Vesta	60
Food	Jala	48
	Jam	29
	Jet	103
	John	128
	Joyce	31
	Jupie	32
	Kyros	45
	Lancia	51
	Lexus	25
	Merkur	59
	Mina	54
	Tabor	38
	Thilo	21
	Vesta	52
PB +	Jala	45
	Jet	45
	John	45
	Joyce	45
	Jupie	45
	Lexus	45
	Manuka	45
	Merkur	45
	Mina	45
	Tabor	45
	Thilo	45
	Vesta	45
PB-	Jala	45
	Jam	45
	Jandira	45
	Jet	60
	John	47
	Joyce	45
	Jupie	45
	Kyros	42
	Lancia	45
	Lexus	51
	Manuka	45
	Merkur	45
	Mina	45
	Tabor	54
	Thilo	45
	Vesta	48
Teasing	Jala	42
	Jam	46
	Jet	47
	John	30
	Joyce	46
	Jupie	60
	Kyros	44
	Lancia	40
	Lexus	44
	Manuka	47
	Mina	59
	Tabor	55
	Thilo	49
	Vesta	60

S.3 Individual information about the thermal reaction and the state of piloerection

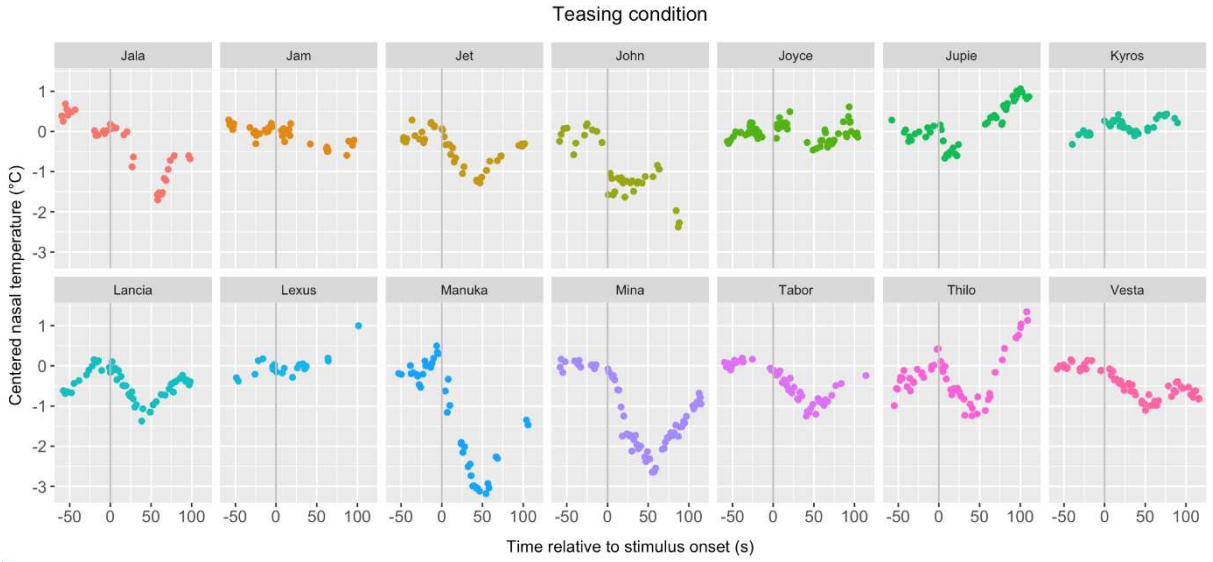
Table S.3 | Information about the thermal reaction and the extent of piloerection for each individual in the five different conditions. The results of the Mann-Whitney U tests show whether there was a significant temperature change from pre to post within a single individual.

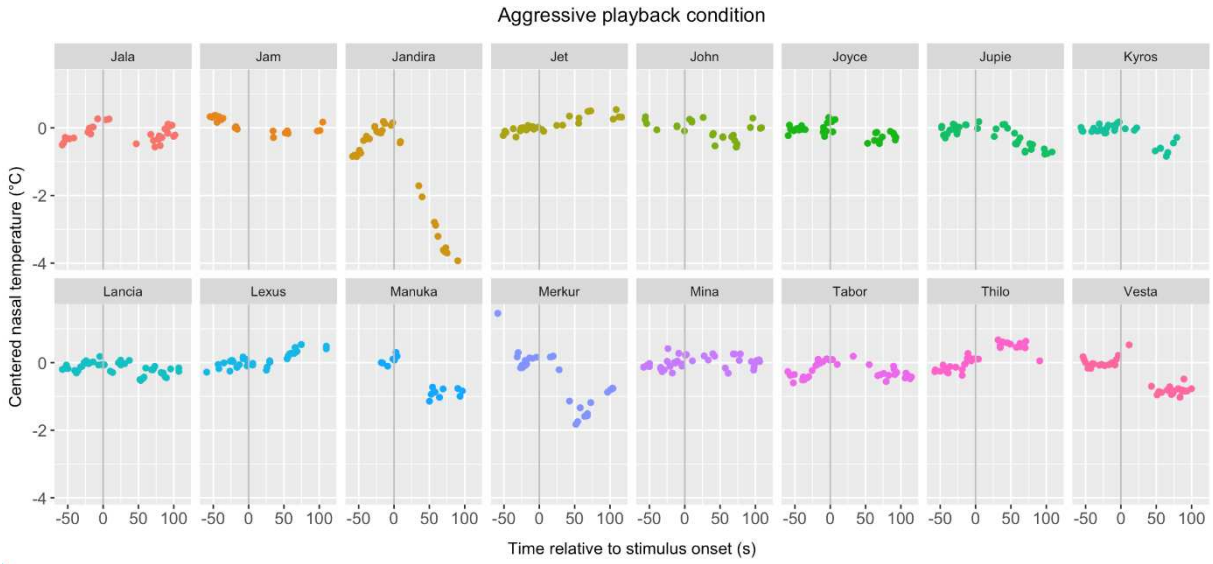
Sex	Subject	Nasal temperature change*	*p*-value of Mann-Whitney U test	Direction of temperature change	Extent of piloerection
Control					
F	Jaja	-0,796	0.000	Decrease	0
	Jala	-0,097	0.065	None	0
	Joyce	-0,626	0.006	Decrease	0,167
	Jupie	-0,491	0.012	Decrease	0
	Lancia	0,23	0.004	Increase	-0,800
	Mina	0,103	0.351	None	0,180
	Vesta	0,454	0.002	Increase	0
M	Jet	0,137	0.025	Increase	0
	Kyros	0,576	0.000	Increase	-0,250
	Lexus	0,196	0.352	None	0,100
	Merkur	0,325	0.001	Increase	0
	Tabor	0,035	0.384	None	0
	Thilo	0,168	0.004	Increase	0
Food					
F	Jala	0,329	0.001	Increase	0,140
	Joyce	0,124	0.142	None	0
	Jupie	0,009	0.659	None	0
	Lancia	0,018	0.645	None	0
	Mina	0,008	1	None	0
	Vesta	-0,147	0.043	Decrease	-0,400
M	Jam	0,37	0.100	Increase	0
	Jet	-0,057	0.775	None	0
	John	0,437	0.008	Increase	0
	Kyros	-0,138	0.315	None	0
	Lexus	0,422	0.011	Increase	0,200
	Merkur	0,169	0.045	Increase	0
	Tabor	0,252	0.010	Increase	0
	Thilo	0,155	0.020	Increase	0
Playback of food calls					
F	Jala	-0,63	0.000	Decrease	0
	Joyce	0,166	0.003	Increase	-0,375
	Jupie	0,094	0.556	None	0
	Manuka	-0,02	0.864	None	0,143

	Mina	-0,334	0.010	Decrease	0,222
	Vesta	-0,246	0.008	Decrease	0,222
M	Jet	0,246	0.012	Increase	-0,250
	John	0,25	0.117	None	0,571
	Lexus	-0,026	0.786	None	0
	Merkur	0,412	0.000	Increase	NA
	Tabor	-0,126	0.330	None	0
	Thilo	0,646	0.000	Increase	0
Playback of aggressive vocalizations					
F	Jala	-0,4	0.010	Decrease	0,111
	Jandira	-3,345	0.000	Decrease	1
	Joyce	-0,315	0.005	Decrease	0
	Jupie	-0,48	0.001	Decrease	0,800
	Lancia	-0,309	0.000	Decrease	0,333
	Manuka	-0,908	0.006	Decrease	1,444
	Mina	0,021	0.913	None	0
	Vesta	-0,856	0.000	Decrease	1,111
M	Jam	-0,134	0.100	None	0,905
	Jet	0,496	0.030	Increase	0
	John	-0,376	0.012	Decrease	0
	Kyros	-0,651	0.000	Decrease	1,625
	Lexus	0,318	0.000	Increase	-1,114
	Merkur	-1,537	0.000	Decrease	0,889
	Tabor	-0,399	0.002	Decrease	1,125
	Thilo	0,519	0.000	Increase	0
Teasing					
F	Jala	-1,468	0.001	Decrease	0,500
	Joyce	-0,266	0.000	Decrease	0
	Jupie	0,454	0.000	Increase	0
	Lancia	-0,892	0.000	Decrease	0,857
	Manuka	-2,744	0.000	Decrease	0,250
	Mina	-1,831	0.001	Decrease	0,077
	Vesta	-0,705	0.000	Decrease	1,133
M	Jam	-0,452	0.001	Decrease	0
	Jet	-0,824	0.000	Decrease	0,900
	John	-1,256	0.001	Decrease	1
	Kyros	0,271	0.002	Increase	1,714
	Lexus	0,158	0.222	None	0,500
	Tabor	-0,755	0.017	Decrease	0,857
	Thilo	-0,571	0.050	Decrease	0,500

* Average nasal temperature of subphase 5

857
858
859


Teasing condition
Jala
Jam
Jet
John
Joyce
Jupie
Kyros
Lancia
Lexus
Manuka
Mina
Tabor
Thilo
Vesta
Centered nasal temperature (°C)
Time relative to stimulus onset (s)


Aggressive playback condition
Jala
Jam
Jandira
Jet
John
Joyce
Jupie
Kyros
Lancia
Lexus
Manuka
Merkur
Mina
Tabor
Thilo
Vesta
Centered nasal temperature (°C)
Time relative to stimulus onset (s)

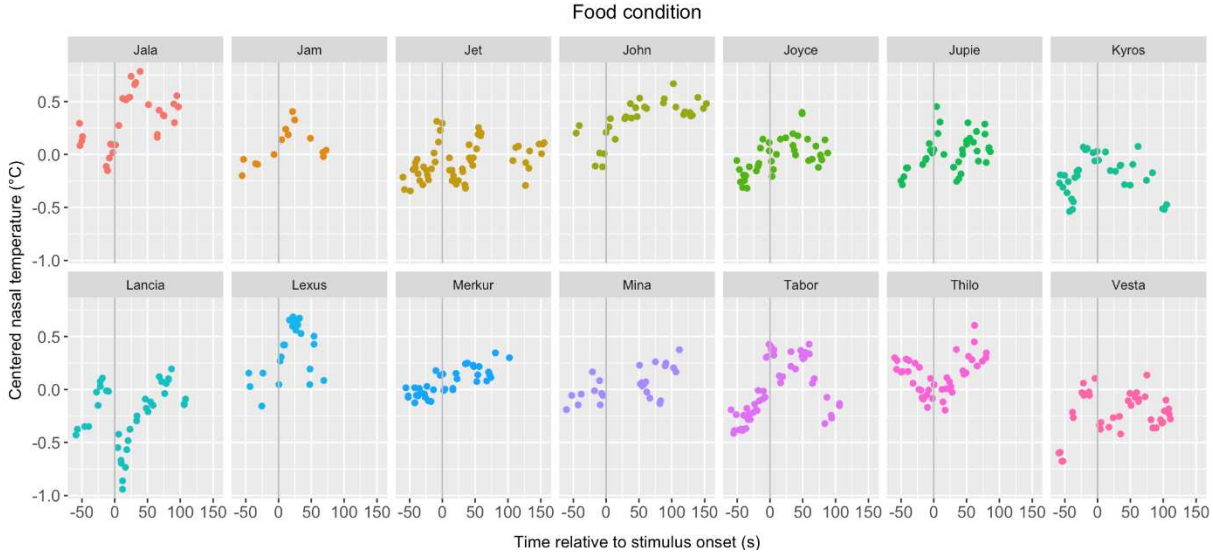

Food condition
Jala
Jam
Jet
John
Joyce
Jupie
Kyros
Lancia
Lexus
Merkur
Mina
Tabor
Thilo
Vesta
Centered nasal temperature (°C)
Time relative to stimulus onset (s)

Figure S.1 | Nasal temperature course over time for each individual in each condition. The stimulus onset at time 0 is indicated by a vertical line.

S.4 Analysis of nasal temperature during baseline

Over all the male monkeys there was a significant increase in nasal temperature in the two positive and the control condition. To test the hypothesis that this pattern could be an artefact of an insufficient habituation to the experimental set-up, we examined nasal temperature during the baseline only in more detail. We were interested in whether males showed a stronger increase in nasal temperature during the baseline (60 seconds before the stimulus started) supporting the hypothsis that they were still habituating. Therefore, we conducted a LMM on the correlation coefficient of the relationship between time and nasal temperature in the baseline of each session. We set sex as fixed effect and subject as random effect.

According to this analysis, sex was not a significant predictor of the correlation coefficient ($\chi^2(2) = 0.901$, $p = 0.637$). This means that males did not show a significantly steeper increase in nasal temperature during the baseline what speaks against the hypothesis that males needed longer to habituate.

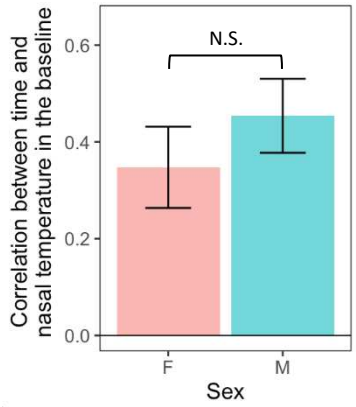


Figure S.2 | Indiviudal correlation coefficients between time and nasal temperature during the baseline, split up by males and females.